Linear Regression Analysis

For Final Project ECO619

You are required to follow these steps for doing and presenting Linear Regression Analysis.

Step # 01: Selection of Variables

In this step, you are required to identify the specific factors (independent variables) which affect the (dependent variable). For the purpose of identification, you must study research articles relevant to your chosen topic. (*You are also advised to seek guidance from the instructor before finalizing the variables).*

After the selection, you will provide brief introduction of the selected variables (dependent as well as independent variables) along with the sources from where the data has been collected. It is highly recommended to collect data from World Development Indicators (WDI) at least for the recent FIVE years or more and test for factors affecting the dependent variable. Also find evidence that out of the selected independent variables which variable is/are significant in determining the dependent variable by using Multiple Linear Regression Analysis.

This step should be presented in the following way

Dependent Variable	**Independent Variables**	**Year Under Consideration**	**Analysis Technique**	**Data Analysis Software**	**Data Collection Source**
		20XX – 20YY	Regression	MS-Excel/ E-Views	WDI
*Choose at least 3 independent variables for regression analysis.					

Step # 02: Model Specification

In this step you will be required to present your model in mathematical equation form.

For example,

$$\mathbf{ER = \alpha + \beta_1 INF + \beta_2 INT + \beta_3 TRO + \varepsilon}$$

Where:

ER = Exchange Rate
INF = Inflation Rate
INT = Interest Rate
TRO = Trade Openness
ε = Error/disturbance/residual

You can also do regression one by one by selecting each independent variable separately in each model to avoid the problem of multi-collinearity. Search articles related to your chosen topic to see the methodology.

Depending to the nature of the variables and data, you can also take the model in logarithm form to make the same units of all variables. In that case, the model can be written as:

$$\mathbf{Log(ER) = \alpha + \beta_1 log(INF) + \beta_2 log(INT) + \beta_3 log(TRO) + \varepsilon}$$

After taking log, the coefficients will be interpreted as elasticities in percentage.

Step # 03: Data Analysis

In this step, you are required to show the output results of the linear regression. Procedure of how to run a Linear Regression in Excel is given in later part of the document.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.5173425
R Square	0.2676432
Adjusted R Square	0.0235243
Standard Error	2.217208
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	10.779481	5.3897409	1.0963644	0.3927968
Residual	6	29.496073	4.916012		
Total	8	40.275555			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95%*	*Upper 95%*
Intercept	1.6181742	1.9731876	0.8200812	0.4435207	-3.2100412	6.4463904	-3.210041	6.4460468
INF	0.2159459	0.3262453	0.6619129	0.5326100	-0.5823475	1.0142395	-0.582347	1.0142335
TRO	0.3276435	0.3317540	0.9876097	0.361481	-0.4841294	1.1394165	-0.484129	1.1394165

Step # 04: Interpretation of the Results

In this step you are required to determine the significant factors that have an impact on dependent variable (Exchange Rate) and also identify their impact level. So, these findings should be based on the interpretation of:

1. Intercept Coefficient
2. Slope Coefficients along with direction of relationship (being positive/negative)
3. R square
4. Model Significance (as per **Prob (F-statistic)**)

1. **a** = 1.618 (Intercept Coefficient)
2. $\mathbf{b_1}$ = 0.216 (Coefficient of Inflation Rate)
   $\mathbf{b_2}$ = 0.327 (Coefficient of Trade Openness)

   So our regression equation is ER = 1.618 + 0.216(INF) + 0.327(TRO)

   $\mathbf{b_1}$ = 0.216 means that one percent increase in inflation rate brings 0.216 percent increase in exchange rate, on average.

   Note: In the above mentioned way, you are required to interpret all the slope coefficients.

   Also in the regression statistics output gives the goodness of fit measure i.e. R square.

3. **R square** = 0.267 (This means 26.7% of variation in exchange rate is determined by inflation rate and trade openness and the rest is unexplained due to the factors/variables not considered/included in the study).

4. **Model Significance (as per probability of F-statistics)** = 1.096 (0.392) means that model/results are not significant as p-value of F '0.392' is greater than 0.05 (significance level). So we can infer that there is no signification relationship among the dependent and independent variables.

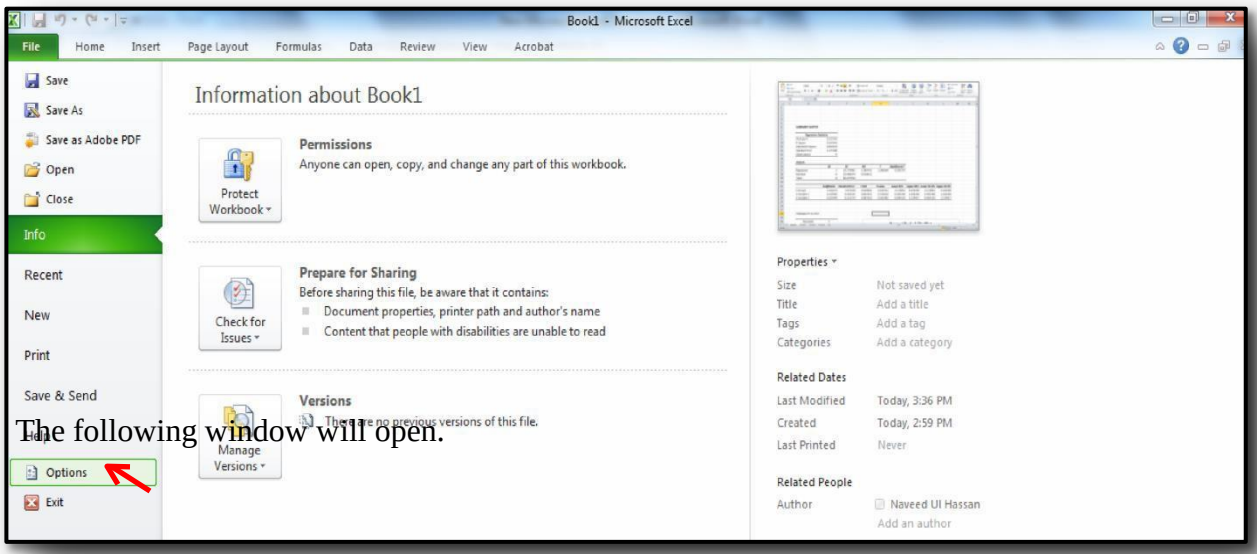
Important Notes

- Regression analysis based on the data older or less than recent five years will **NOT** be accepted.
- You are required to submit you the **Raw Data (excel file)** which contains the data of variables along with your project report.

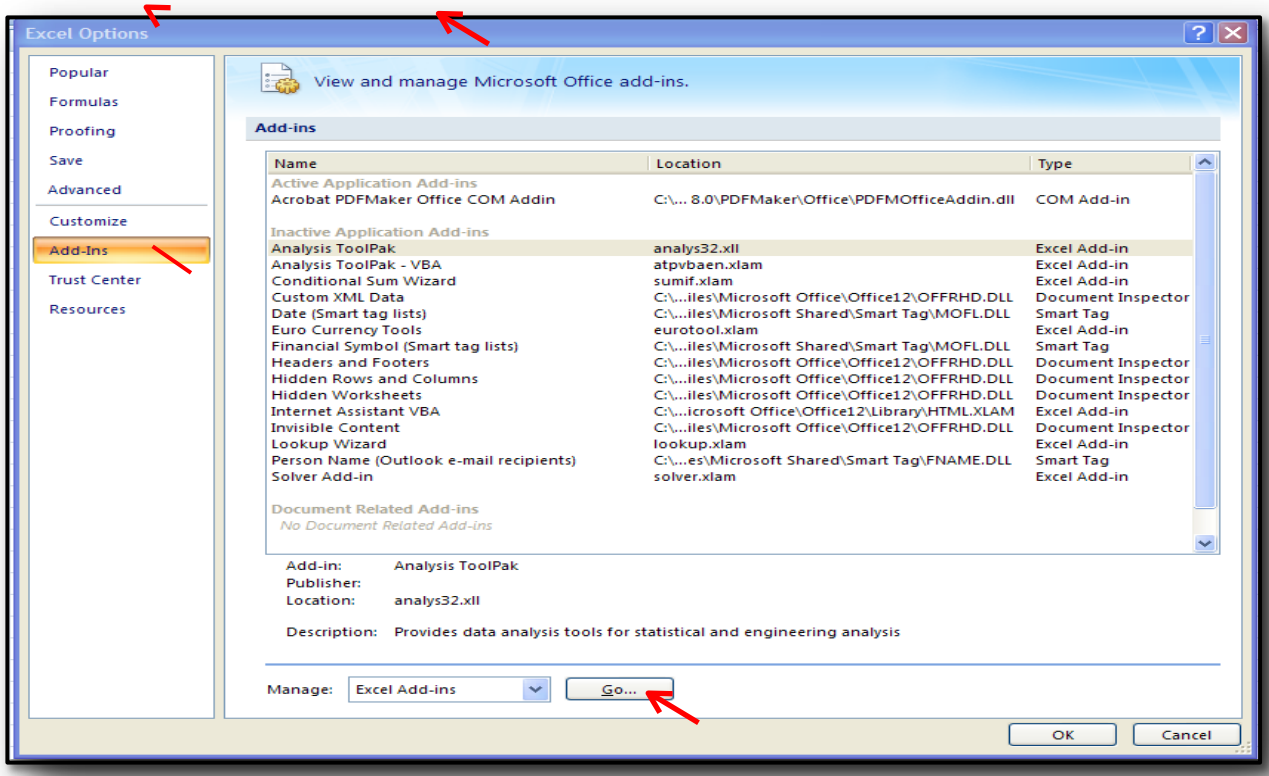
Linear Regression using Excel

Installing the Analysis ToolPak

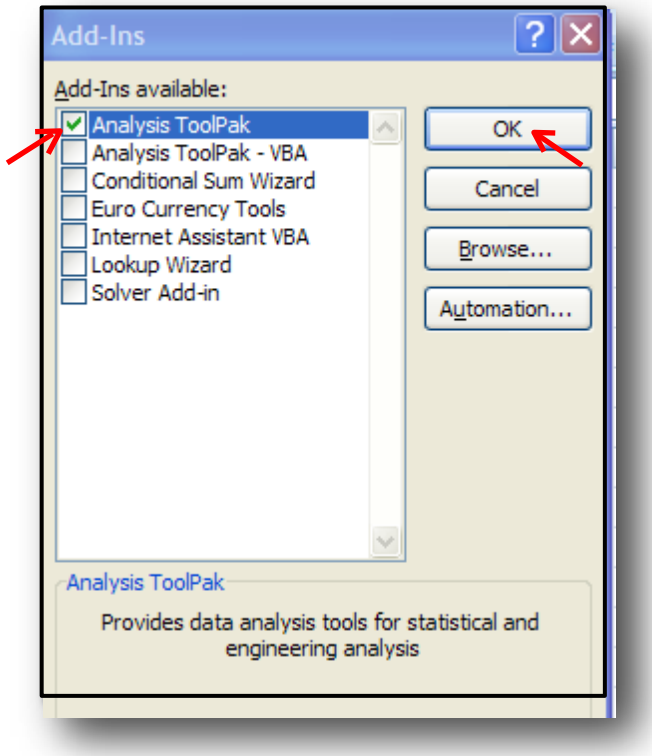
1. Open Microsoft Excel, click the File menu and select **Options**



The following window will open.



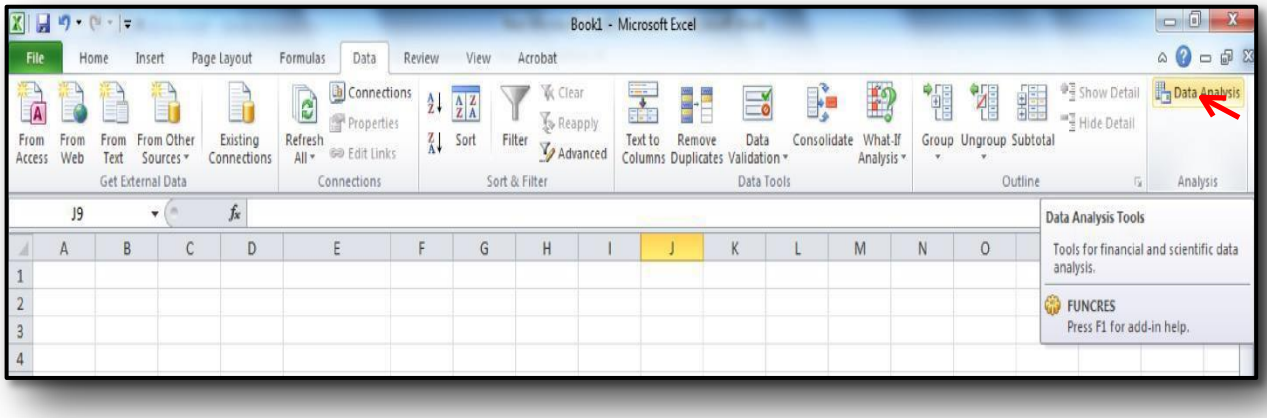
2. In the Microsoft Office button, go to excel options to click Add-ins
3. In the Add-Ins box, select Analysis ToolPak and click Go...



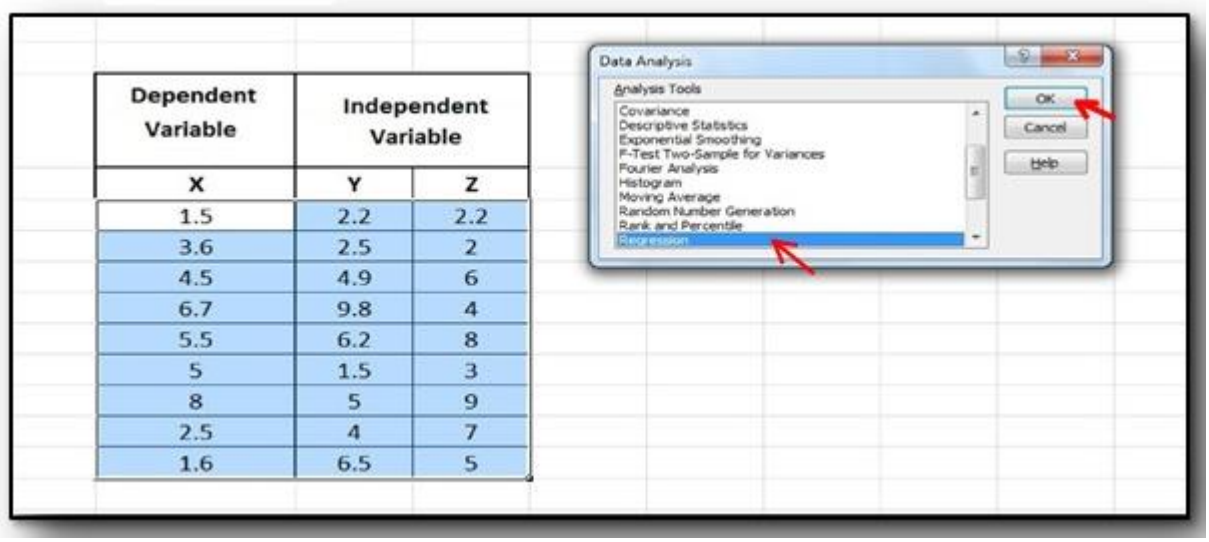
4. In the Add-Ins available box, check the Analysis ToolPak and then click OK.
5. If Analysis ToolPak is not listed in the Add-Ins available box, click Browse to locate it.

Linear Regression using the Data Analysis Add-In

After installation of Analysis ToolPak you will find Data analysis button on the right corner of data menu. Click this button



Following window will open.



1. In the data analysis tool select the regression and then click Ok.

Suppose we want to determine whether X is a function of Y and Z.

$$\mathbf{X_i = a + b_1 Y_i + b_2 Z_i + error_i}$$

Where;

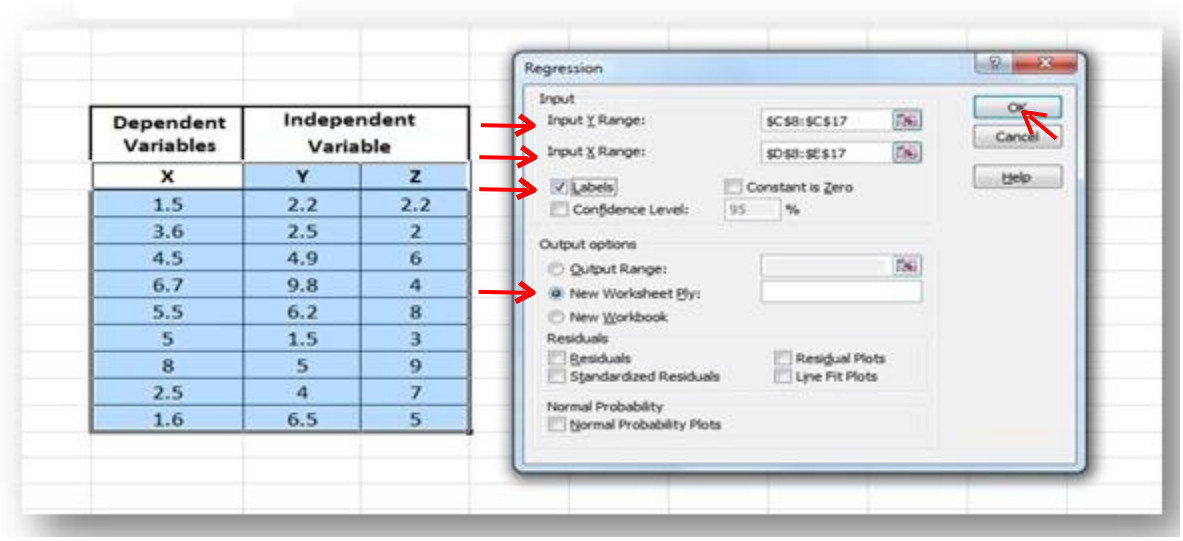
X_i = value of X for observation i

a = mean value of X (dependent variable) when Y and Z (independent Variable) is zero (intercept coefficient)

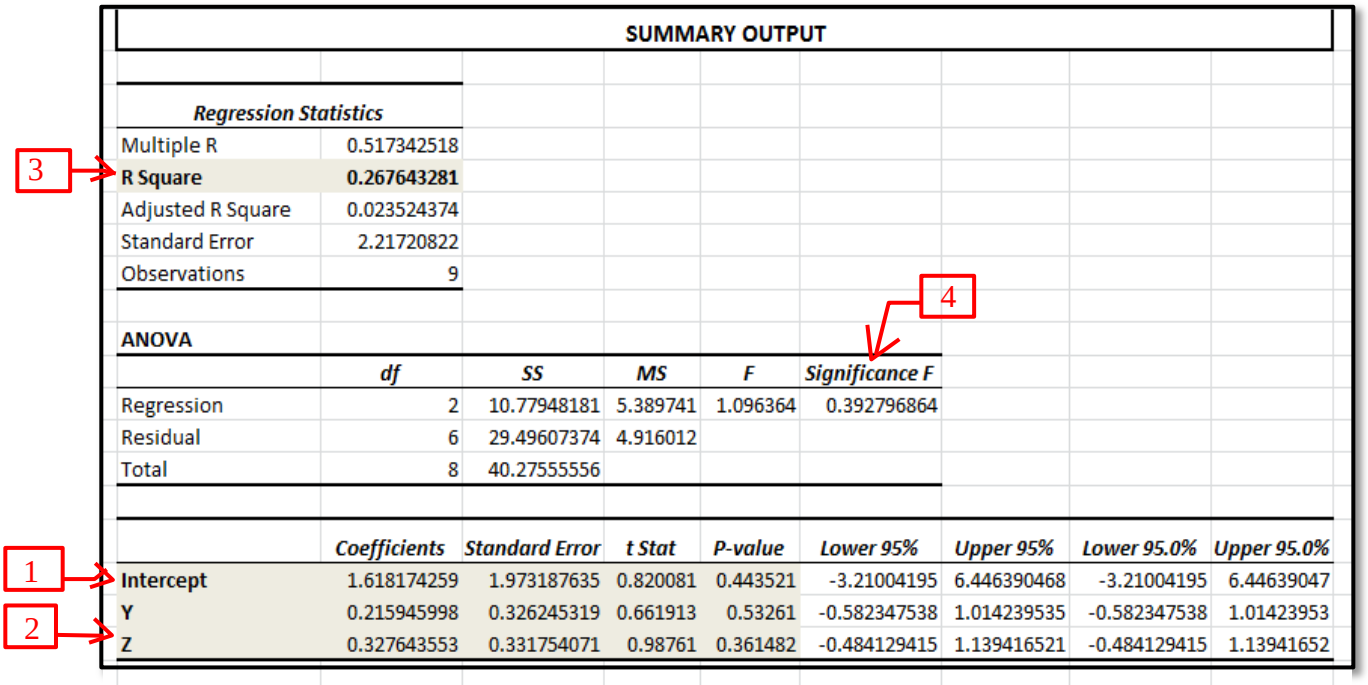
b = average change in X given a one unit change in Y and Z, i.e. (slope of Y and Z)

b_1 = value of Y for observation i

b_2 = value of Z for observation i



2. After clicking the regression, a regression box appears.
3. Select the Inputs for Y range (Dependent Variable) in first box.
4. Select the Inputs for X range (Independent Variable) in 2nd box.
4. Select the place where you want your output
5. Check the Labels and then click Ok



SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.517342518
R Square	**0.267643281**
Adjusted R Square	0.023524374
Standard Error	2.21720822
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	10.77948181	5.389741	1.096364	0.392796864
Residual	6	29.49607374	4.916012		
Total	8	40.27555556			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	1.618174259	1.973187635	0.820081	0.443521	-3.21004195	6.446390468	-3.21004195	6.44639047
Y	0.215945998	0.326245319	0.661913	0.53261	-0.582347538	1.014239535	-0.582347538	1.01423953
Z	0.327643553	0.331754071	0.98761	0.361482	-0.484129415	1.139416521	-0.484129415	1.13941652

The output is given in the coefficients column in the last set of output

1. Intercept Coefficient
2. Slope Coefficients
3. R square
4. Model Significance (as per **Prob (F-statistic)**)